

(19) 世界知的所有権機関  
国際事務局(43) 国際公開日  
2001年8月9日 (09.08.2001)

PCT

(10) 国際公開番号  
WO 01/58096 A1

(51) 国際特許分類<sup>7</sup>: H04L 12/56

(21) 国際出願番号: PCT/JP01/00663

(22) 国際出願日: 2001年1月31日 (31.01.2001)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:  
09/496,706 2000年2月2日 (02.02.2000) US

(71) 出願人: 三菱電機株式会社 (MITSUBISHI DENKI KABUSHIKI KAISHA) [JP/JP]; 〒100-8310 東京都千代田区丸の内二丁目2番3号 Tokyo (JP).

(72) 発明者: ヴェトロ アンソニー (VETRO, Anthony); 10314 ニューヨーク州 ステートン・アイランド レジス・ドライブ 113 New York (US). ディヴァカラン アジェイ (DIVAKARAN, Ajay); 07834 ニュージャー

ジー州 デンヴィル コブ・ロード 20 New Jersey (US). スンハイファン (SUN, Huifang); 08572 ニュージャー ジー州 クランベリー キングレット・ドライブ・サウス 61 New Jersey (US).

(74) 代理人: 曾我道照, 外(SOGA, Michiteru et al.); 〒100-0005 東京都千代田区丸の内三丁目1番1号 国際ビルディング8階 曾我特許事務所 Tokyo (JP).

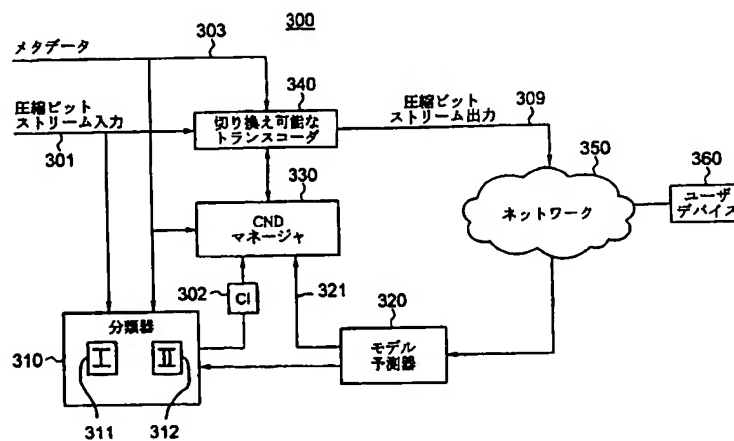
(81) 指定国 (国内): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) 指定国 (広域): ARIPO 特許 (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), ユーラシア特許 (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ特許 (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,

[続葉有]

(54) Title: METHOD AND APPARATUS FOR DISTRIBUTING COMPRESSED BIT STREAM TO USER DEVICE OVER NETWORK

(54) 発明の名称: 圧縮ビットストリームをネットワークを通してユーザデバイスに配信するための装置及び方法



303...METADATA  
301...COMPRESSED BIT STREAM INPUT  
340...SWITCHABLE TRANSCODER  
309...COMPRESSED BIT STREAM OUTPUT

330...CND MANAGER  
350...NETWORK  
360...USER DEVICE  
320...MODEL PREDICTOR  
310...CLASSIFIER

656542546

WO 01/58096 A1

(57) Abstract: A multimedia distribution system for distributing compressed bit streams to user devices through a network comprises a transcoder and a manager. The transcoder processes bit streams using any one of a plurality of conversion modes. A manager selects one from a plurality of conversion modes according to semantic contents and network characteristics of bit streams. The system also comprises a content classifier for determining the content characteristics, and a model predictor for determining network characteristics and user device characteristics. The manager includes an integrator that generates an optimum rate-quality function used to select a particular conversion model for a predetermined bit rate available in the network.

[続葉有]



LU, MC, NL, PT, SE, TR), OAPI 特許 (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

---

(57) 要約:

ネットワークを通してユーザデバイスに圧縮ビットストリームを配信するためのマルチメディア配信システムは、トランスコーダおよびマネージャを有する。トランスコーダは、複数の変換モードのいずれか1つを用いてビットストリームに対して動作するように構成される。マネージャは、ビットストリームの意味論的コンテンツおよびネットワーク特性に応じて、複数の変換モードの特定の1つを選択するように構成される。システムはまた、コンテンツ特性を決定するためのコンテンツ分類器、ならびにネットワーク特性およびユーザデバイス特性を決定するためのモデル予測器を有する。マネージャのインテグレータは、ネットワークの所定の利用可能なビットレートに対して特定の変換モデルを選択するために用いられる最適なレート-品質関数を生成する。

## 明 細 書

圧縮ビットストリームをネットワークを通してユーザデバイスに  
配信するための装置及び方法

5

技術分野

この発明は、一般に、情報配信システムに関し、特に、圧縮ビットストリームとして符号化された情報をネットワークで利用可能なビットレートおよびビットストリームの意味論的コンテンツに適応させる配信システムに関する。

10

背景技術

近年、符号化情報を通信するための数多くの規格が開発されてきた。デジタル画像について、最も良く知られている規格は J P E G である。Pennebacker  
ら「JPEG Still Image Compression Standard」、Van Nostrand Reinhold、19  
15 93を参照のこと。画像シーケンスで最も広範囲に用いられている規格としては、M P E G - 1（動画の格納および取り出し用）、M P E G - 2（デジタルテレビ用）、および H. 2 6 3 が挙げられる。ISO/IEC JTC1 CD11172、MPEG、「Information Technology - Coding of Moving Pictures and Associated Audio  
for Digital Storage Media up to about 1.5 Mbit/s - Part2: Coding of  
20 Moving Pictures Information」、1991年、LeGall、「MPEG: A Video Compression Standard for Multimedia Applications」、Communications of the ACM  
第34巻4号、46～58頁、1991年、ISO/IEC DIS 13818-2、MPEG-2、「Information Technology - Generic Coding of Moving Pictures and Associated Audio  
Information - Part2: Video」、1994年、ITU-T SG XV、DRAFT H.263、「Video  
25 Coding for Low Bitrate Communication」1996年、ITU-T SG XVI、DRAFT13  
H.263+ Q15-A-60 rev.0、「Video Coding for Low Bitrate Communication」、1997年を参照下さい。

これらの規格は、主に画像の空間的圧縮と画像シーケンスの空間的および時間的圧縮を取り扱う比較的低レベルの仕様である。一般的な特徴として、これ

らの規格は、フレーム毎の圧縮を行う。これらの規格によると、広範囲な応用で高い圧縮比を成し遂げることができる。

MPEG-4（マルチメディア応用）などのさらに新しい画像コーディング規格（「Information Technology -- Generic coding of audio/visual objects」ISO/IEC FDIS 14496-2（MPEG4 Visual）、1998年11月を参照）では、任意の形状のオブジェクトを個別の画像オブジェクトプレーン（VOP：Video Object Planes）として符号化および復号化することができる。オブジェクトは、視覚データ、音声データ、自然データ、合成データ、基本データ、複合データ、またはその組み合わせであり得る。画像オブジェクトは、複合オブジェクトまたは「シーン：scenes」を形成するように組み立てられる。

新しく浮上しつつあるMPEG-4規格は、自然および合成マテリアルが統合され、アクセスが普遍的であるインタラクティブ画像などのマルチメディア応用を可能にすることを意図している。MPEG-4は、コンテンツを基礎とした相互作用を可能にする。例えば、移動する像またはオブジェクトを1つの画像から他の画像に「カットアンドペースト：cut-and-paste」したい場合がある。このタイプの応用では、マルチメディアコンテンツ内のオブジェクトは、何らかのタイプの分割プロセスを通して識別されてきたと想定される。例えば、1999年6月4日付けで出願された、Lin等による米国特許出願第09/326,750号「Method for Ordering Image Spaces to Search for Object Surfaces」を参照下さい。

画像伝送のコンテキストにおいて、これらの圧縮規格は、ネットワークで必要とされる帯域幅（利用可能なビットレート）の量を減少させるために必要である。ネットワークは、無線チャネルまたはインターネットを表し得る。いずれにせよ、ネットワークは、限定された容量を有し、そのリソースに対するコンテンツは、コンテンツを送信しなければならない場合に解決されなければならない。

長年の間、デバイスにコンテンツをロバストに送信し、コンテンツの品質を利用可能なネットワークリソースに適応させることが可能なアーキテクチャおよびプロセスに対して多大な努力がなされてきた。コンテンツがすでに符号化

されている場合、ネットワークを通してストリームが送信される前に、すでに圧縮されたビットストリームをさらに変換し、例えば、利用可能なビットレートを減少させなければならないことがある。

図1に示すように、通常、スケーリングは、トランスコーダ100によって成し遂げられ得る。力任せの場合、トランスコーダは、復号化器110および符号化器120を有する。圧縮された入力ビットストリーム101は、入力レート $R_{in}$ で完全に復号化され、次いで、新しい出力レート $R_{out}$ 102で符号化され、出力ビットストリーム103が生成される。通常、出力レートは入力レートよりも低い。しかし、実際には、復号化されたビットストリームの符号化は非常に複雑であるため、トランスコーダにおける完全な復号化および完全な符号化はなされない。

MPEG-2 トランスコーディングに関する初期の研究については、Sun等によって、「Architectures for MPEG compressed bitstream scaling」、IEEE Transactions on Circuits and Systems for Video Technology、1996年4月において公開されている。この文献では、複雑度およびアーキテクチャが異なる、レートを低下させるための4つの方法が示されている。

図2は、例示的な方法を示している。このアーキテクチャでは、画像ビットストリームは、一部のみが復号化される。具体的には、入力ビットストリーム201のマクロブロックは、可変長復号化(VLD)210される。入力ビットストリームはまた遅延され220、逆量子化(IQ)230され、離散コサイン変換(DCT)係数を形成する。所望の出力ビットレートが与えられると、部分的に復号化されたデータは分析され240、250において新しい集合の量子化器がDCTブロックに適用される。次に、これらの再量子化されたブロックは、可変長コード化され(VLC)260、より低いレートの新しい出力ビットストリーム203が形成され得る。この方式は、図1に示す方式よりもはるかに簡単である。なぜなら、動きベクトルが再使用され、逆DCTオペレーションを必要としないからである。

Assuncao等によるさらに最近の研究では、「A frequency domain video transcoder for dynamic bit-rate reduction of MPEG-2 bitstream」、IEEE Tra

nsactions on Circuits and Systems for Video Technology、953～957頁、1998年12月において、同じタスクに関する簡略化されたアーキテクチャが記載されている。ここでは、ドリフト補償のために周波数ドメインにおいて動作する動き補償（MC）ループが用いられる。近似マトリクスは、周波数ドメイン内のMCブロックを迅速に計算するために引き出される。トランスコーディングのための最良の量子化器スケールを計算するためにラグランジュ最適化が用いられる。

Sorial等「Joint transcoding of multiple MPEG video bitstreams」、Proceedings of the International Symposium on Circuits and Systems、Can 1999年による他の研究は、多重MPEG-2ビットストリームを共にトランスコード化する方法を提示している。Vetro等によって1999年10月1日付けで提出された、米国特許出願第09/410,552号「Estimating Rate-Distortion Characteristics of Binary Shape Data」も参照下さい。

従来の圧縮規格によると、テクスチャ情報を符号化するために割り付けられたビットの数は、定量化パラメータ（QP：quantization parameter）によって制御される。上記の文献は、元のビットストリームに含まれる情報に基づいてQPを変化させることによってテクスチャビットのレートを低下させるという点では同様である。効率的なインプリメンテーションでは、情報は、通常、圧縮されたドメイン内で直接抽出され、DCTブロックのマクロブロックまたは残留エネルギーのモーションに関連する測定を含み得る。このタイプの分析は、ビット割り付け分析器において見出され得る。

場合によっては、ビットストリームは予め処理され得るが、トランスコーダがリアルタイムで動作することも重要である。従って、ビットストリームに関する大幅な処理遅延は許容できない。例えば、トランスコーダがフレーム群から情報を抽出し、次いで、このルックアヘッド情報に基づいてコンテンツをトランスコード化することは実現可能ではない。これは、生放送またはビデオ会議では作用しない。より良好なビット割り付けにより、品質に関してはより良好なトランスコーディングを成し遂げることは可能であるが、このようなリアルタイム応用のインプリメンテーションは実用的ではない。

従来のトランスコーディング方法では、ビットレートを低下させる能力が限定されていることに留意することも重要である。換言すると、出力画像のQPのみが変化する場合、どのくらいのレート低下が可能であるかには限界がある。低下は、考慮されているビットストリームに依存して限定される。QPから

5 最大値に変更すると、通常、ビットストリームのコンテンツは大幅に低下する。空間的品質を低下させるものとしては、この他に、時間的品質の低下、即ちフレームを落したりまたはスキップすることが挙げられる。あまりに多くのフレームをスキップすることも、品質を大幅に低下させることになる。空間的品質および時間的品質の低下が共に考慮される場合、トランスコーダは、空間的

10 品質対時間的品質におけるトレードオフに直面する。しかしながら、そのような空間的および時間的な減少でさえも、オリジナル画像ビットストリームによって伝えられるコンテンツを破壊することなくそのターゲットレートに応じることはトランスコーダにとって難しい。

この結果、トランスコーダは、利用可能なビットレートの低下に適應するように、ビットストリームに含まれる情報を送信する何らかの代替手段を見出さなければならない。

15

### 発明の開示

この発明は、圧縮ビットストリームをネットワークを通してユーザデバイス

20 に配信するためのマルチメディア配信システムを提供する。このシステムは、トランスコーダおよびマネージャを有する。トランスコーダは、複数の変換モードのいずれか1つを用いて、ビットストリームに対して動作するように構成される。マネージャは、ビットストリームの意味論的コンテンツおよびネットワーク特性に応じて、複数の変換モードの特定の1つを選択するように構成さ

25 れる。システムはまた、コンテンツ特性を決定するためのコンテンツ分類器、ならびにネットワーク特性およびユーザデバイス特性を決定するためのモデル予測器を有する。マネージャのインテグレータは、ネットワークの所定の利用可能なビットレートに対して特定の変換モデルを選択するために用いられる最適なレート-品質関数を生成する。

### 図面の簡単な説明

- 図1は、従来のトランスコーダのブロック図、  
図2は、従来の部分復号化器／符号化器のブロック図、  
5 図3は、この発明による適応可能なビットストリーム配信システムのブロック図、  
図4は、この発明による適応可能なトランスコーダおよびトランスコーダマネージャのブロック図、  
図5は、図4のトランスコーダおよびマネージャによって用いられ得るトランスコーディング関数のグラフである。  
10

### 発明を実施するための最良の形態

#### [ビットストリームの品質：Quality of Bitstream]

- 圧縮された入力ビットストリームを、ターゲットレート（即ち、ネットワークで利用可能なビットレート（ABR：available bit rate））で圧縮された出力ビットストリームに変換することが可能なビットストリーム配信システムについて記載する。通常、ターゲットレートは、オリジナルレートよりも小さい。換言すると、トランスコーダのタスクは、通常、ネットワークリソースにおける制約のために、ビットストリームをさらに圧縮することである。  
15  
我々のシステムの新規性は、従来のトランスコーダの欠点、即ち、特にリアルタイム応用におけるレート変換の制限を克服することができるということである。従来のトランスコーディング技術は十分にレートを低下させることができるものの、コンテンツの品質は通常、激しく低下する。大抵の場合、ビットレートが低下したビットストリームで伝達される情報は全く失われてしまう。  
20  
従来、ビットストリーム「品質」は、入力ビットストリームと出力ビットストリームとの間のビット毎の差として測定されてきた。

ビットストリームのコンテンツの品質を維持すると共に、ターゲットレートを成し遂げることができるトランスコーディング技術について記載する。



## [連続変換 : Continuous Conversion]

従来のフレームを基礎とするトランスコーディング技術は、連続変換と定義され得る。従来の技術は、空間対時間的品質において最良のトレードオフを連続して維持することを試みるため、出力は常に、入力シーケンスを最良に表示するフレームのシーケンスである。レート上の制約を満たすため、特定のフレームがスキップされるとき、スキップされたフレーム内に含まれる情報は考慮されない。十分なフレームがスキップされると、受信されるビットストリームは、ユーザにとって意味のないものになるか、良くても満足のいくものではない。

10

## [品質歪みメトリクス : Quality Distortion Metrics]

従来の連続変換トランスコーダは、空間および時間的品質におけるトレードオフに関して、レート歪みの観点で最適な決定を下す。このようなトランスコーダでは、歪みは、通常、ノイズ比に対するピーク信号 (P SNR) などの従来の任意の歪みメトリクスとしてとられる。このような変換では、歪みは、ビットストリームのコンテンツがどのくらい良好に伝達されているかの測定ではなく、むしろ元の入力ビットストリームと再構築された出力ビットストリームとの間のビット間の差 (即ち、品質) である。

## 20 [ビットストリームの忠実度 : Fidelity of Bitstream]

この発明に係る低ビットレート制約下でビットシーケンスをトランスコード化するための1つの実施の形態において、少数のフレームを有するビットストリームのコンテンツについて要約する。この方法において、我々は品質に焦点を当てた従来の歪みメトリクスを用いない。むしろ、「忠実度 (fidelity)」と呼ばれる新しい基準を採用する。忠実度は、コンテンツの意味論を考慮する。意味論は、ビットまたは画素を意味せず、むしろビットで表される人間にとって意味のある概念、例えば、単語、音、画像オブジェクト、シーンなどを意味する。

忠実度は多くの方法で定義され得る。しかし、忠実度は、ここで定義するよ

うに、従来の量的な品質、例えば、ビット間の差には関連しない。むしろ、忠実度は、1つのフレームまたは任意の数のフレームが元の画像シーケンスに含まれる情報を伝達する程度、即ち、伝達される情報のコンテンツまたはより高レベルな意味を測定し、生ビットを測定するのではない。

5

[離散要約トランスコーダ: Discrete-Summary Transcoder]

忠実度は、従来の歪みメトリクスに比べ、より主観的または意味論的な測定である。しかし、我々のシステムでは、忠実度は、従来のものではないトランスコーダの性能を評価するための有用な測定である。一実施の形態による我々の

10 のトランスコーダの出力は、ビットシーケンス全体の要約を試みる比較的高い品質のフレームの限定された集合であるため、我々は、このタイプのトランスコーダを「離散要約トランスコーダ」と呼ぶ。

例えば、低ビットレートでは、我々は、画像を表示するために少数の高品質なフレームを選択する。このようにして、ビットストリームの意味論上の「意

15 味」が保存される。この離散要約トランスコーダが入力ビットストリームの高レベルな意味論的サンプリングを成し遂げる一方、連続したトランスコーダは、空間および時間ドメインにおいて画素を量的にサンプリングするだけであることに留意されたい。ビットレートがかなり限定されている状況では、我々は、「豊富な (rich)」フレームをサンプリングし、ビットストリーム内の符号

20 化されたコンテンツの忠実度を保存する。

豊富なフレームを選択的にサンプリングするため、ビットストリームにおける1つの局面、即ち動画(性能)を失い得る。連続変換トランスコーダのレート歪み性能が激しく低下するかまたはターゲットレートを満足することができない場合にのみ、好ましくは離散要約トランスコーディングをとる。これらの

25 条件下では、従来の連続変換トランスコーダは、滑らかな動画(性能)を失う。なぜなら、フレームレートが非常に低いため、情報配信レートは画像がぎくしゃくし(jerkyと呼ぶ現象をきたし)ユーザに不快感を与えるからである。

従来の連続変換トランスコーディングに対する離散要約トランスコーディングの主な利点は、厳しいレート制約下にある連続変換トランスコーダが、情報

が豊富なフレームを落とすのに対して、離散要約トランスコーダは情報が豊富なフレームを選択するよう試みることである。

所定の状況に対してどのトランスコーダが最良であるかを制御するために、コンテンツネットワークデバイス（CND）マネージャについて記載する。CNDマネージャの目的は、どのトランスコーダを用いるかを選択することである。選択は、コンテンツ、ネットワーク、ユーザデバイス特性から得られるデータに基づいている。我々はまた、「オフライン」モードにおけるこれらのデバイス特性をシミュレートし、ビットストリームを変化させ、後に配信することができる。

10

[適応可能なビットストリーム配信システム：Adaptable Bitstream Delivery System]

図3に示すように、適応可能なビットストリーム配信システム300は、4つの主な構成要素、即ち、コンテンツ分類器310、モデル予測器320、コンテンツネットワークデバイスマネージャ330、および切り換え可能なトランスコーダ340を有する。

システム300の目標は、圧縮されたビットストリーム301を情報コンテンツと共にネットワーク350を通してユーザデバイス360に配信することである。ビットストリームのコンテンツは、視覚データ、音声データ、テキストデータ、自然データ、合成データ、基本データ、複合データ、またはその組み合わせであり得る。ネットワークは、無線、パケット切り換え型であるか、または予測不可能な動作特性を有する他のネットワークであり得る。ユーザデバイスは、画像受信機、固定型または移動型無線受信機、またはビットストリームの品質受信を困難にし得る内部リソース制約を有する同様の他のユーザデバイスであり得る。

利点として、システムは、ビットストリームがネットワークおよびユーザデバイス特性を満足するようにさらに圧縮される必要があるときでさえ、コンテンツの意味論的忠実度を維持する。

入力圧縮ビットストリームは、トランスコーダおよびコンテンツ分類器に向

けられる。トランスコーダは、最終的には、ネットワークを通してユーザデバイスに向けられる出力圧縮ビットストリーム 309 のレートを低下させ得る。

- コンテンツ分類器 310 は、マネージャ用の入力ビットストリームからコンテンツ情報 (CI) 302 を抽出する。コンテンツ分類器の主な機能は、モーション活動、画像変化情報およびテクスチャなどのコンテンツ特性の意味論的特徴を、コンテンツネットワークマネージャにおいてレート-品質トレードオフをなすために用いられるパラメータの集合にマップすることである。このマッピング機能を補助するために、コンテンツ分類器はまた、メタデータ情報 303 を受け得る。メタデータは、低レベルおよび高レベルであり得る。メタデータの例としては、新しく浮上しつつある MPEG-7 規格によって特定される記述子および記述方式が含まれる。

- このアーキテクチャでは、モデル予測器 320 は、ネットワーク 350 のダイナミクスに関し、恐らくはユーザデバイス 360 の特性を制約するリアルタイムフィードバック 321 を提供する。例えば、予測器は、ネットワーク輻輳および利用可能なビットレート (ABR) を報告する。予測器はまた、ネットワーク内でのパケット損失比に関するフィードバックを受信し、翻訳する。予測器は、現在のネットワーク状態、および長期ネットワーク予測 321 を見積もる。典型的には、ユーザデバイスは、リソースが限定され得る。例えば、処理能力、メモリ、および表示の制約である。例えば、ユーザデバイスがセルラー電話である場合、その表示は、テキスト情報または低レゾリューション画像、またはさらに悪い場合には、単なる音声に制約され得る。これらの特性はまた、トランスコーディング様式を選択にも影響を与え得る。

- メタデータ 303 の受信に加えて、マネージャ 330 はまた、コンテンツ分類器 310 およびモデル予測器 320 の両方から入力を受信する。CND は、切り換え可能なトランスコーダ 340 に対して最適なトランスコーディング方策が決定されるように、これらの 2 つの情報源からの出力データを組み合わせる。

[コンテンツ分類器 : Content Classifier]

パターン分析および認識の分野では、分類は、様々なレベルの画像から特徴を抽出することによって成し遂げることができる。例えば、プログラム特徴、ショット特徴、フレーム特徴、およびフレーム内のサブ領域の特徴である。特徴自体は、高性能な変換または簡単ローカルオペレータを用いて抽出され得る

5   。特徴がどのように抽出されるかに関係なく、寸法Nの特徴空間が与えられると、各パターンは、この特徴空間内のポイントとして表示され得る。

種々の異なるトレーニングパターンを入力としてこの抽出プロセスにかけ、その結果を特徴空間内でプロットすることはよく行われている。特徴集合およびトレーニングパターンが適切である場合、「クラス」と呼ばれるいくつかの

10   ポイントのクラスが観察される。これらのクラスによって、異なるパターンを識別し、同様のパターンをグループ化することができ、観察されたクラス間の境界を画定することができる。通常、クラス間の境界は、分類ミスのためいくらかのコストは免れないが、全体としてのエラーを最小限にするよう試みる。

適切なクラスが識別され、クラス間の適切な境界線が引かれると、ビットストリーム内の新しいパターンをすばやく分類することができる。問題によっては、これは、ニューラルネットワークまたはサポートベクトル機械などの他の公知の分類技術（Cristianiniら、「An Introduction to Support Vector Mac

15   hines（および他のかゝるをベースとした学習方法）」、Cambridge University Press、2000を参照）を用いて成し遂げることができる。

コンテンツ分類器310は、2ステージ（IおよびII）において動作する。第1に、より高レベルな意味論が推論され得るようにビットストリームコンテンツを分類し、第2に、分類されたコンテンツをネットワークおよびユーザデバイス特性に適応する。

25   第1のステージ（I）311では、従来の技術を用いて圧縮ビットストリームから多数の低レベルの特徴、例えば、モーション及びテクスチャを抽出する。また、MPEG-7記述子および記述方式などのメタデータ303にもアクセスすることができる。メタデータが利用可能である場合、圧縮ビットストリームに対してはそれほどの作用は必要ない。この第1のステージの最終結果と

して、予め決定されたコンテンツ特徴の集合が意味論的クラスまたは高レベルのメタデータの限定された集合にマップされる。さらに、各意味論的クラス内で、コーディングの複雑度（即ち、複雑度は、意味論的クラスおよびネットワーク特性、ならびに恐らくはデバイス特性を条件とする）に基づいて区別する

5     。

このコンテンツの高レベルの理解は、コンテンツ情報（CI）302としてCNDマネージャ330に渡される。CI302は、この発明に係る切り換え可能なトランスコードの潜在的な性能を部分的に特徴づける。

上記の分類は、コンテンツ理解、および最終的には離散要約トランスコーディングの観点から有用であるが、中間ステージの結果としても有用である。本質的には、分類の第2ステージII 312への入力として作用する新しい集合のクラスを有する。分類の第2ステージでは、意味論的なクラスをネットワークの特徴およびデバイス特性にマップする。これらの特徴によって、システムがトランスコーディング方策を開発するのを補助するレート品質関数の特性を決定するのを助ける。換言すると、特定の意味論的クラスが、オブジェクトの移動または画像の変化のために、バーストデータによって特徴づけられる可能性がある場合、このことは、どのくらいのリソースをネットワークが提供しなければならぬかを見積もる際に考慮されなければならない。第3のステージ313については、他の実施の形態に関して以下に記載する。

20

[コンテンツネットワークデバイスマネージャ：Content-Network-Device Manager]

コンテンツネットワークデバイス（CND）マネージャ330およびトランスコード340は、図4にさらに詳細に示される。CNDマネージャは、離散連続制御431およびコンテンツネットワークデバイス（CND）インテグレート432を有する。トランスコード340は、複数のトランスコード441～443を有する。

制御431は、スイッチ450を用いて、例えば、離散要約トランスコード441、連続変換トランスコード442、または何らかの他のトランスコード

4 4 3で、入力圧縮ビットストリーム3 0 1がどのようにトランスコード化されるべきかを決定する機能を有する。ネットワークコンテンツマネージャはまた、トランスコードに対するターゲットレートに動的に適応し、ネットワークおよびユーザデバイスの特性を制約するリソースを考慮する。これらの2つの  
5 非常に重要な項目は、制御4 3 1によって決定される。

どのように制御が最適な選択決定をなすかをより良く理解するために、図5は、複数のレート-品質関数をレート5 0 1および品質5 0 2のスケールに関してグラフで示している。連続変換トランスコード4 4 2の1つのレート-品質関数は、凸関数5 0 3によって示される。離散要約トランスコード4 4 1に  
10 対するレート-品質曲線は、一次関数5 0 4によって表される。他のトランスコードは異なる関数を有し得る。

これらの曲線が単に例示を目的として描かれたものであることに留意されたい。特定のトランスコードに対する関数の真の形式は、コンテンツ、コンテンツがどのように分類されたか、および恐らくはネットワークおよび特性を制約  
15 するデバイスの現在の状態に応じて変化し得る。明らかに、低ビットレートでは、上記の理由のために、連続変換トランスコードは品質が急速に劣化する。最適な品質関数5 0 5は太字で示される。この関数は、所定のビットレートおよびユーザデバイスに対して成し遂げられ得る最適な品質を最良にモデル化する。

20 我々は、レート=T 5 0 6においてトランスコーディング技術ではクロスオーバーが発生することに留意する。Tよりも大きなレートについては、連続変換トランスコードを用い、Tよりも小さいレートについては、離散要約トランスコードを用いることが最良である。言うまでもなく、クロスオーバーポイントは、コンテンツおよびネットワーク特性が変化するにつれて動的に変化する。

25 上述したように、連続変換トランスコードは、通常、P S N Rなどの従来の歪みメトリクスを想定する。このような測定は、我々の離散要約トランスコードには適用されないため、従来の歪みメトリクスを「忠実度」の測定にマップすることはより妥当である。忠実度は、コンテンツがどのくらい良好に意味論的に要約されるかを測定し、量的なビット間の差は測定しない。同じ品質メト

リクスが与えられると、最適なトランスコーディング方策を決定する際の矛盾を避ける。

## 5 [コンテンツネットワークデバイスインテグレータ：Content-Network-Device Integrator]

図4を再び参照すると、CNDインテグレータ432は、コンテンツ分類器310からのコンテンツ情報302と、モデル予測器からのネットワークデバイス予測321とを共に組み合わせるCNDマネージャの部分である。図5に示されるレート品質関数、または他の同様の最適化関数として表されるモデルを生成するのはマネージャのこの部分である。最適な動作モデル321を形成するために、CNDインテグレータは、コンテンツ分類器からのマッピングCIおよび切り換え可能なトランスコーダ340から出力されるビットレートフィードバック351を調べる。この情報を用いて、インテグレータは、特定のモデルパラメータを有する最適なモデリング関数505を選択する。レートフィードバック351は、パラメータを動的に改良するために用いられる。インテグレータが、選択されたモデルが最適でないことを発見する場合、インテグレータは、レート品質関数を動的に切り換える決定をすることができる。また、インテグレータは、異なるオブジェクトまたは異なるビットストリームに対していくつかの関数を追跡し、関数を個別にまたは一緒に考慮し得る。

20

## [ネットワーク予測の影響：Impact of Network Predictions]

ネットワーク予測321は、最適曲線505の特定部分を一方向または他方向に変調することによってこれらの特性関数に影響を与え得る。例えば、より高いビットレートが利用できる場合、最も注意を払う必要がある。ネットワークモデルによって、特定時間で多数のビットを消費することができるが、長期の影響によって、急速に混雑し易いことが分かるため、我々のシステムは、抑制して、より低いレートで動作を続けることを選択し得る。このようにして、利用可能なビットレートの突然の低下に関する問題を回避する。これらのタイプの特性は、トランスコーダの曲線を変調することによって引き起こされ得る



。

## [デバイス制約の影響 : Impact of Device Constrains]

また、デバイス特性を考慮する必要がある。移動型デバイスは、固定型デバイスとは異なる動作特性を有する。例えば、ドップラースプレッドは、利用可能なビットレートが高いほど性能が低下し得る。従って、より低いビットレートを  
5     選択しなければならない。デバイスは、トランスコーダに影響を与え得る限定された処理、格納、および表示能力を有し得る。例えば、画像をオーディオのみのデバイスに配信することは意味がない。事実、切り換え可能なトランスコーダは、スピーチからテキストまたはデータからスピーチへの変換などを行  
10    う他のトランスコーダ 4 4 3 を有し得る。重要な点は、この発明の切り換え可能なトランスコーダがビットストリームコンテンツの意味論および目的地デバイスを考慮するのに対して、多くの従来のトランスコーダは利用可能なビットレートを考慮するだけであるということである。

15

## [トランスコーダ : Transcoder]

トランスコーダの種々のタイプを満たす技術は従来技術において簡単に見つけられることができるので、我々はトランスコーダの操作を詳しく述べない。例えば、以下の任意の米国特許第 5, 9 9 1, 7 1 6 号 (スピーチのタンデム  
20    コーディングを防止するトランスコーダ)、第 5, 9 4 0, 1 3 0 号 (抽出された動き補償データをバイパス転送する画像トランスコーダ)、第 5, 7 6 8, 2 7 8 号 (N : 1 トランスコーダ)、第 5, 7 6 4, 2 9 8 号 (緩和された内部復号化器 / コーダインタフェースフレームジッタ要件を有するデジタルデータトランスコーダ)、第 5, 5 2 6, 3 9 7 号 (切り換えトランスコーダ)  
25    、第 5, 3 3 4, 9 7 7 号 (異なるビット数がコード変換に用いられる ADPCM トランスコーダ)、または他の同様の特許を参照のこと。

本実施の形態で強調されるのは、ビットストリームの意味論的コンテンツの最良の配信を提供するトランスコーディング方策の動的選択を可能にすることであり、実際のトランスコーディングがどのように実行されるかではない。

この発明を好ましい実施の形態の例により記載したが、他の様々な適応および改変は、この発明の精神および範囲内であり得ることを理解されたい。従って、添付の請求の範囲の目的は、このような変更および改変のすべてをこの発明の真の精神および範囲内でカバーすることである。

## 請 求 の 範 囲

1. 圧縮ビットストリームをネットワークを通してユーザデバイスに配信するための装置であって、

5 複数の変換モードのいずれか1つを用いて、前記ビットストリームに対して動作するように構成されるトランスコーダと、

前記ビットストリームの意味論的コンテンツおよびネットワーク特性に応じて、前記複数の変換モードの特定の1つを選択するように構成されるマネージャとを備えた装置。

10

2. 前記マネージャは、ユーザデバイス特性に応じて、前記複数の変換モードの特定の1つを選択する請求項1に記載の装置。

3. 前記ビットストリームを受信するように接続された分類器と、前記  
15 ネットワークおよびデバイス特性を受信するように接続されたモデル予測器とをさらに備えた請求項2に記載の装置。

4. 前記ビットストリームのコンテンツは、視覚データ、音声データ、  
テキストデータ、自然データ、合成データ、基本データ、複合データ、および  
20 その組み合わせのいずれかであるように選択される請求項1に記載の装置。

5. 前記分類器は、前記ビットストリームからコンテンツ情報を抽出し、前記コンテンツ特性の意味論的特徴の地図をつくる請求項3に記載の装置。

25 6. 前記分類器は、前記ビットストリームのメタデータを受信し、前記コンテンツ特性の意味論的特徴の地図をつくる請求項5に記載の装置

7. 前記ネットワーク特性は、ネットワーク輻輳、利用可能なビットレート、およびパケット損失比を含む請求項1に記載の装置。

8. デバイス特性は、処理能力、メモリ、および表示の制約を含む請求項 2 に記載の装置。

9. 前記分類器は、前記ビットストリームの意味論的コンテンツを分類する請求項 3 に記載の装置。

10. 前記分類されたコンテンツは、前記ネットワーク特性、前記デバイス特性、および前記トランスコーダからのレートフィードバックに従って分類される請求項 3 に記載の装置。

11. 前記マネージャは、制御およびインテグレータを有する請求項 1 に記載の装置。

12. 前記トランスコーダは、それぞれが前記複数の変換モードの 1 つにおいて動作する複数のトランスコーダを有する請求項 11 に記載の装置。

13. 前記複数の異なるトランスコーダは、離散要約トランスコーダおよび連続変換トランスコーダを有する請求項 12 に記載の装置。

14. 前記マネージャは、前記コンテンツおよびネットワーク特性から複数のレート-品質関数を生成する請求項 12 に記載の装置。

15. 前記制御は、前記ネットワークの特定のビットレートに対する前記複数のレート-品質関数を最大にすることに応じて、前記複数の変換モードの特定のモードの選択を行う請求項 14 に記載の装置。

16. 前記選択は動的である請求項 15 に記載の装置。

17. 圧縮ビットストリームをネットワークを通してユーザデバイスに

配信するための方法であって、

前記ビットストリームの意味論的コンテンツを分類するステップと、

前記ネットワークの動作を予測するステップと、

前記意味論的コンテンツおよび前記ネットワークの前記予測された動作に応

- 5    じて、トランスコードの複数の変換モードの1つを選択するステップとを含む方法。

18.    前記ユーザデバイス特性に応じて、前記トランスコードの前記複数の変換モードの1つを選択するステップをさらに含む請求項17に記載の方法。
- 10

図1

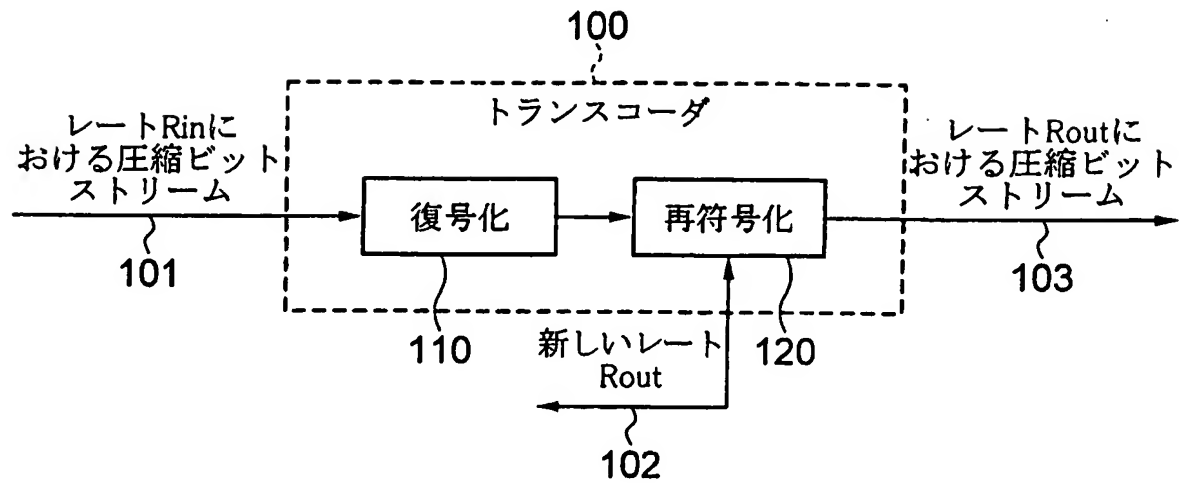


図2

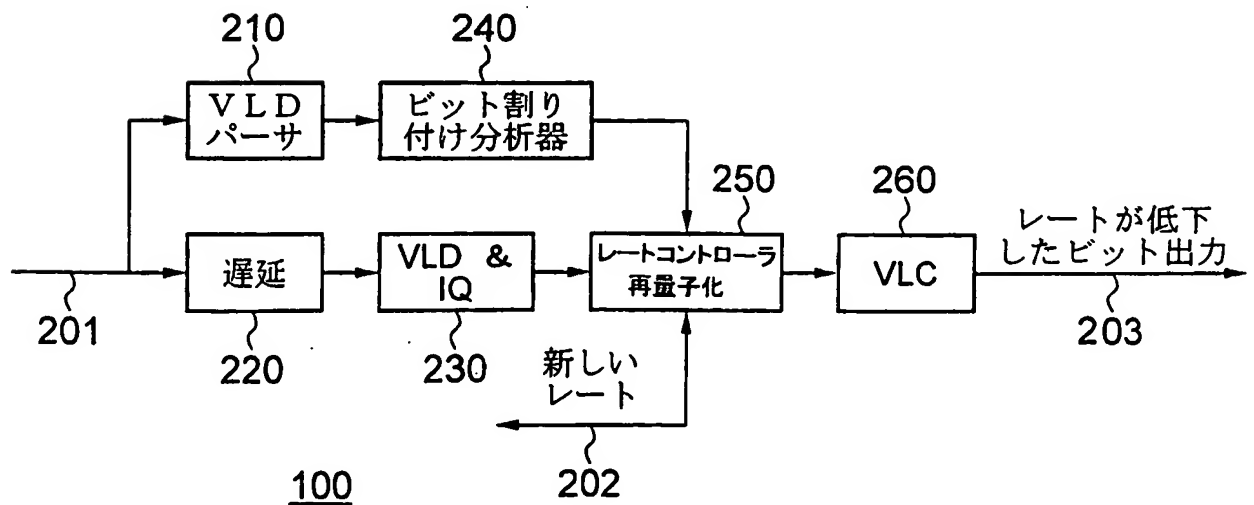


図3

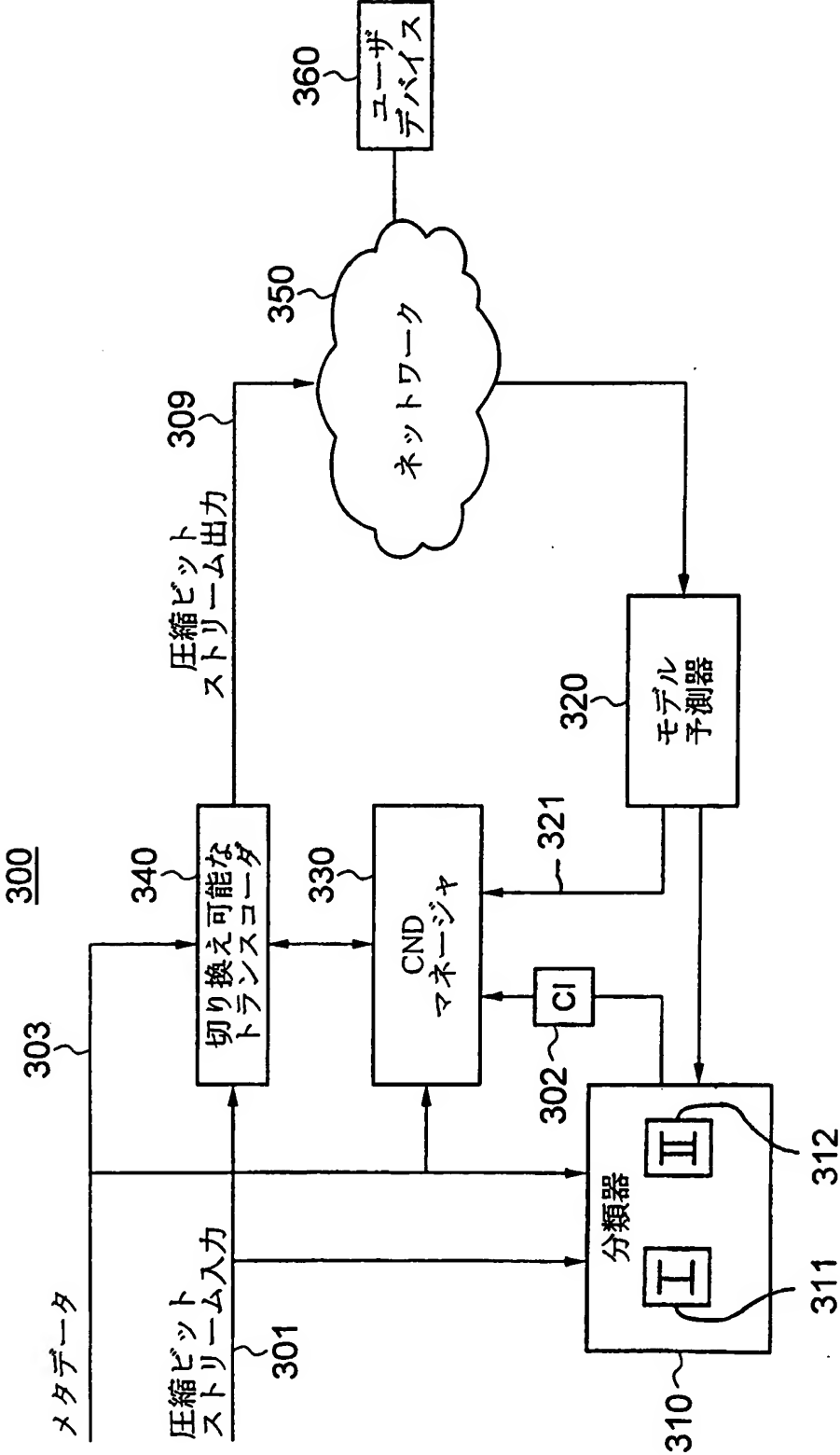


図4

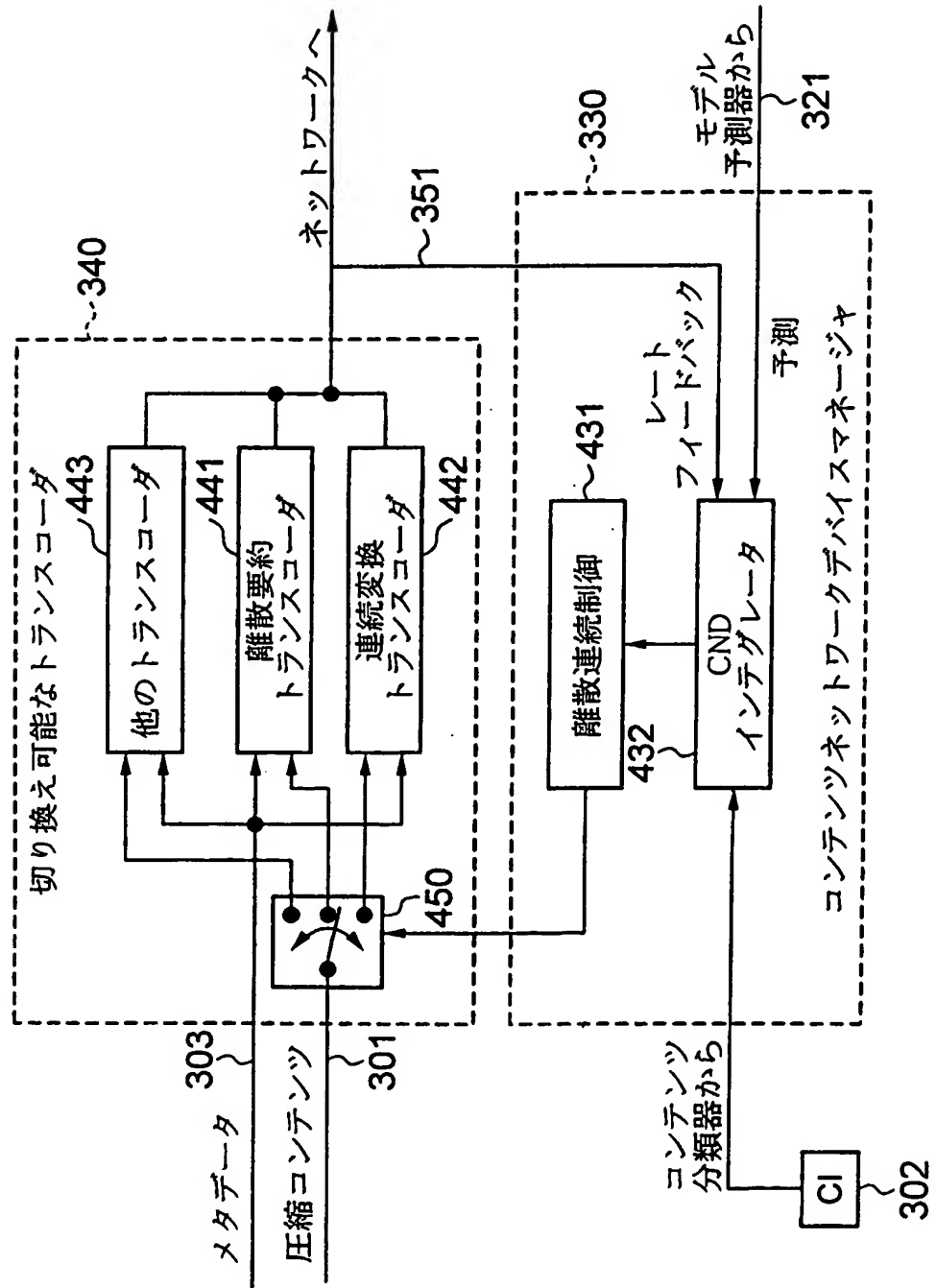
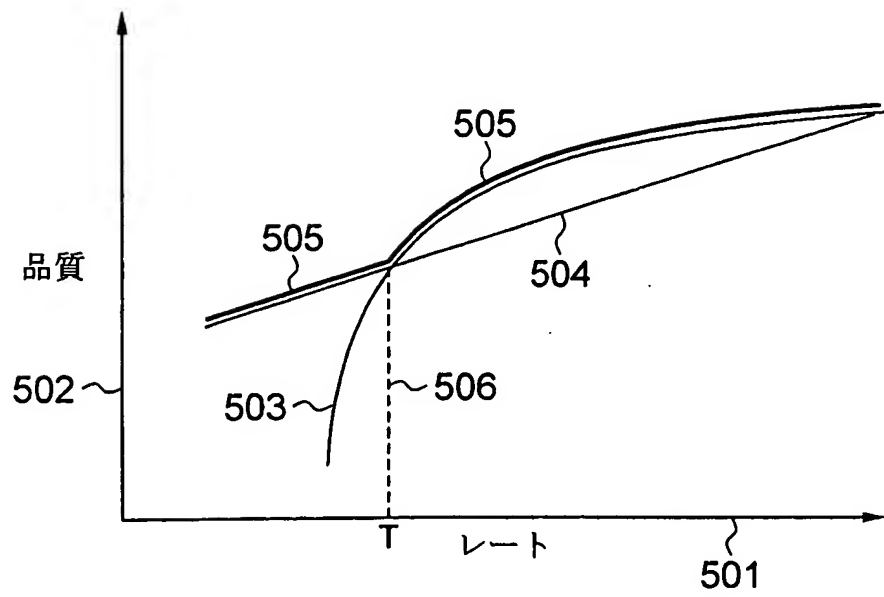




図5



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/00663

A. CLASSIFICATION OF SUBJECT MATTER  
Int.Cl.<sup>7</sup> H04L12/56

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
Int.Cl.<sup>7</sup> H04L12/56

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Technical report of IEICE : IN , SSE , CQ , CS  
the Institute of Electronics, Information Communication Engineers, General Meeting,  
the Institute of Electronics, Information Communication Engineers, Communication Society Meeting

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP, 1-180185, A (NEC Corporation), 18 July, 1989 (18.07.89) (Family: none)	1-18
X	JP, 1-138837, A (Fujitsu Limited), 31 May, 1989 (31.05.89) (Family: none)	1-18
X	JP, 11-298528, A (Fujitsu Limited), 29 October, 1999 (29.10.99) (Family: none)	1-18
X	JP, 4-77150, A (NEC Corporation), 11 March, 1992 (11.03.92) (Family: none)	1-18
A	US, 5140417, A (Matsushita Electric Co., Ltd.), 18 August, 1992 (18.08.92) & JP, 3-22780, A	1-18
A	US, 5604731, A (Lucent Technologies Inc.), 18 February, 1997 (18.02.97) & EP, 739114, A & CA, 2172318, A & JP, 8-298512, A	1-18

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>
--	---

Date of the actual completion of the international search  
27 April, 2001 (27.04.01)

Date of mailing of the international search report  
15 May, 2001 (15.05.01)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP01/00663

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, 5559798, A (Lucent Technologies Inc.), 24 September, 1996 (24.09.96) & EP, 741474, A & CA, 2172319, A & JP, 8-198511, A	1-18
A	JP, 8-32598, A (Hitachi, Ltd.), 02 February, 1996 (02.02.96) (Family: none)	1-18
A	JP, 4-316294, A (Matsushita Electric Ind. Co., Ltd.), 06 November, 1992 (06.11.92) (Family: none)	1-18
A	JP, 10-200531, A (Nippon Telegr. & Teleph. Corp. <NTT>), 31 July, 1998 (31.07.98) (Family: none)	1-18
A	JP, 5-56064, A (Fujitsu Limited), 05 March, 1993 (05.03.93) (Family: none)	1-18
A	JP, 8-25192, A (Toshiba Corporation), 27 September, 1996 (27.09.96) (Family: none)	1-18
A	TECHNICAL REPORT OF IEICE, SSE97-66 (28.08.97), Kentarou FUKUDA et al., "The relationship between QoS prarameters and required bandwidth in MPEG-2 video"	1-18
E,A	JP, 2000-196599, A (NEC Corporation), 14 July, 2000 (14.07.00) (Family: none)	1-18

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int Cl<sup>7</sup> H04L12/56

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int Cl<sup>7</sup> H04L12/56

最小限資料以外の資料で調査を行った分野に含まれるもの

Technical report of IEICE : IN, SSE, CQ, CS

電子情報通信学会総合大会

電子情報通信学会通信ソサイエティ大会

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP, 1-180185, A (日本電気株式会社) 18. 7月 1989 (18. 07. 89) (ファミリー無し)	1-18
X	JP, 1-138837, A (富士通株式会社) 31. 5月 1989 (31. 05. 89) (ファミリー無し)	1-18
X	JP, 11-298528, A (富士通株式会社) 29. 10月 1999 (29. 10. 99) (ファミリー無し)	1-18

☒ C欄の続きにも文献が列举されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&amp;」同一パテントファミリー文献

国際調査を完了した日

27. 04. 01

国際調査報告の発送日

15.05.01

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

吉田 隆之

電話番号 03-3581-1101 内線 3502

5X

9077



C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP, 4-77150, A(日本電気株式会社)11. 3月1992(11. 03. 92) (ファミリー無し)	1-18
A	US, 5140417, A(Matsushita Electric Co., Ltd.)18. 8月1992 (18. 08. 92) & JP, 3-22780, A	1-18
A	US, 5604731, A(Lucent Technologies Inc.)18. 2月1997(18. 02. 97) & EP, 739114, A & CA, 2172318, A & JP, 8-298512, A	1-18
A	US, 5559798, A(Lucent Technologies Inc.)24. 9月1996(24. 09. 96) & EP, 741474, A & CA, 2172319, A & JP, 8-198511, A	1-18
A	JP, 8-32598, A(株式会社日立製作所)02. 2月1996(02. 02. 96) (ファミリー無し)	1-18
A	JP, 4-316294, A(松下電器産業株式会社)06. 11月1992(06. 11. 92) (ファミリー無し)	1-18
A	JP, 10-200531, A(日本電信電話株式会社)31. 7月1998(31. 07. 98) (ファミリー無し)	1-18
A	JP, 5-56064, A(富士通株式会社)05. 3月1993(05. 03. 93) (ファミリー無し)	1-18
A	JP, 8-25192, A(株式会社東芝)27. 9月1996(27. 09. 96) (ファミリー無し)	1-18
A	TECHNICAL REPORT OF IEICE, SSE97-66, 28. 8月1997(28. 08. 97), Kentarou FUKUDA et al, "The relationship between QoS parameters and required bandwidth in MPEG-2 video"	1-18
E, A	JP, 2000-196599, A(日本電気株式会社)14. 7月2000(14. 07. 00) (ファミリー無し)	1-18